JOURNAL OF MULTIVARIATE ANALYSIS 11, 280–288 (1981)

Bayes Minimax Estimation of Multiple Poisson Parameters*

MALAY GHOSH AND AHMAD PARSIAN

Iowa State University

Communicated by A. Cohen

For the p-variate Poisson mean, under the sum of weighted squared error losses, weights being reciprocals of variances, a class of proper Bayes minimax estimates dominating the usual estimate, namely the sample mean is produced. An example is given to illustrate this. The interrelation of our results with those of Clevenson and Zidek is pointed out.

1. INTRODUCTION

Let $\mathbf{X} = (X_1,..., X_p)'$, where $X_1,..., X_p$ are p independent Poisson variables with respective parameters $\theta_1,..., \theta_p$, $\theta_i \in (0, \infty)$ for all $i = 1,..., p$. Denote by $\mathbf{x} = (x_1,..., x_p)$ a realization of $\mathbf{X}$. The problem is the point estimation of $\boldsymbol{\theta} = (\theta_1,..., \theta_p)'$ under losses of the type

$$L_k(\boldsymbol{\theta}, \mathbf{a}) = \sum_{i=1}^{p} \theta_i^{-k}(a_i - \theta_i)^2, \qquad k = 0, 1, 2,..., \tag{1.1}$$

where $\mathbf{a} = (a_1,..., a_p)'$.

The usual estimator of $\boldsymbol{\theta}$ is $\mathbf{X}$. For $p = 1$, under any loss of the type (1.1), X_1 is an admissible minimax estimator of θ_1. However, for any $p \geqslant 2$, $\mathbf{X}$ is no longer an admissible estimator of $\boldsymbol{\theta}$, under any loss of the type (1.1) with $k > 0$ though it continues to be minimax in all these cases except when $k = 2$. The above inadmissibility result was first proved by Clevenson and Zidek [4] for $k = 1$, and later by Tsui and Press [13] for all $k = 1, 2,...$. When $k = 0$, i.e., when the sum of squared error losses is used, the admissibility of $\mathbf{X}$ for $p = 2$ follows from a more general result of Peng [8], while its inadmissibility for $p \geqslant 3$ was proved by Peng [8], Hudson [6] and

Received October 29, 1979; revised June 4, 1980.

AMS 1970 subject classification: 62F10, 62C15, 62H99.

Key words and phrases: p-variate Poisson means, Bayes estimation, admissibility, minimaxity, two stage prior, exponential, type II beta, Dirichlet, monotone likelihood ratio.

* Research supported by NSF Grant MCS-7900274.

0047-259X/81/020280–09$02.00/0

Tsui [12]. These results signify the presence of the Stein effect not only for the normal and other location parameter family of distributions, but also for other types of distributions some of which need not even be continuous. In fact, some of the Poisson inadmissibility results have been extended by Hudson [6] and Tsui [11] to other discrete power series distributions under the sum of squared error losses.

In this note, we consider the loss (1.1) with $k = 1$. In this case Clevenson and Zidek [4] produced a class of minimax estimators dominating $\mathbf{X}$. Also, they were able to find a subclass of such estimators that are Bayes with respect to some proper prior for $p \geqslant 3$ and admissible. Tsui and Press [13] had slightly broadened the minimax class of Clevenson and Zidek [4], but did not investigate any proper Bayes subclass in their bigger class of minimax estimators.

We obtain, in Section 2, a class of proper Bayes estimators dominating $\mathbf{X}$ for $p \geqslant 3$. The interrelation of our results with those of Clevenson and Zidek [4] has been pointed out. In fact, the class of proper Bayes minimax estimators suggested in their Theorem 2.5 turns out to be a subclass of a class of estimators suggested in our example. Brown [3] gives a heuristic reasoning which suggests that the inadmissibility in a p-dimensional Poisson case is analogous to the inadmissibility in a $2p$-dimensional normal case when the loss L_1 is used. It is interesting to note that whereas in the normal case, proper Bayes minimax estimators dominating the usual one exist only for $p \geqslant 5$, under the loss L_1, the class of estimators proposed by us (dominating $\mathbf{X}$) is proper Bayes minimax for $p \geqslant 3$. We have also shown that when $p = 2$, no proper prior meets the requirements of our main theorem leading to generalized Bayes estimators. The question of whether there exists for $p = 2$ any proper Bayes estimators of the type introduced in Section 2 is still open.

We may mention in passing that the determination of the critical dimension of inadmissibility under the sum of weighted squared error losses depends on the nature of weights. This is demonstrated in the work of Berger [1], Brown [3], Tsui [12] and Tsui and Press [13].

2. The Results

Let $X_1, \ldots, X_p$ be independent, $X_i \sim \text{Poisson}(\theta_i)$, $1 \leqslant i \leqslant p$. Suppose that conditional on $T = t$, $\theta_1, \ldots, \theta_p$ are iid exponential with parameter t, i.e., $\boldsymbol{\theta}$ has pdf

$$f(\boldsymbol{\theta} \mid t) = \prod_{i=1}^{p} \{t \exp(-t\theta_i)\} = t^p \exp\left(-t \sum_{1}^{p} \theta_i\right), \tag{2.1}$$

$0 < \theta_i < \infty$ for all $1 \leqslant i \leqslant p$, $0 < t < \infty$. Also, let T have the prior (possibly improper) density $g(t)$ such that the resulting posterior density of $\boldsymbol{\theta}$ given $\mathbf{X} = \mathbf{x}$ is proper. The following lemma provides a convenient expression for the generalized Bayes estimator of $\boldsymbol{\theta}$.

LEMMA 1. *Under the above specification of the prior density for* $\boldsymbol{\theta}$, *the generalized Bayes estimator of* $\boldsymbol{\theta}$ *is given by* $\{E[(1+T) \mid Z]\}^{-1}\mathbf{X}$, *where* $Z = \sum_1^p X_i$.

Proof. The conditional pdf of $\boldsymbol{\theta}$ given $\mathbf{X} = \mathbf{x}$ and $T = t$ is

$$f(\boldsymbol{\theta} \mid \mathbf{x}, t) \propto \prod_1^p \{\theta_i^{x_i} \exp(-(1+t)\,\theta_i)\}. \tag{2.2}$$

Hence, under the loss (1.1) with $k = 1$, the Bayes estimate of θ_i given $X_i = x_i$ ($x_i = 1, 2, \ldots$) and $T = t$ is

$$\begin{aligned}
&\{E(\theta_i^{-1} \mid x_i, t)\}^{-1} \\
&\quad = \int_0^\infty \theta_i^{x_i} \exp(-(1+t)\,\theta_i)\, d\theta_i \Big/ \int_0^\infty \theta_i^{x_i - 1} \exp(-(1+t)\,\theta_i)\, d\theta_i \\
&\quad = \{\Gamma(x_i + 1)(1+t)^{-(x_i+1)}\}/\{\Gamma(x_i)(1+t)^{-x_i}\} \\
&\quad = (1+t)^{-1} x_i \qquad (1 \leqslant i \leqslant p).
\end{aligned} \tag{2.3}$$

For $x_i = 0$, using (2.2) and the fact that

$$\int_0^\infty \theta_i^{-1}(\theta_i - a)^2 \exp(-(1+t)\,\theta_i)\, d\theta_i \begin{cases} = +\infty & \text{if } a \neq 0 \\ < \infty & \text{if } a = 0, \end{cases} \tag{2.4}$$

it follows that the Bayes estimate of θ_i is 0. Hence, in general, the Bayes estimate of θ_i conditional on $X_i = x_i$ and $T = t$ is $(1+t)^{-1} x_i$ $(1 \leqslant i \leqslant p)$. Hence, for $x_i \neq 0$

$$\begin{aligned}
\delta_i(\mathbf{x}) &= \{E(\theta_i^{-1} \mid \mathbf{x})\}^{-1} = \{E[E(\theta_i^{-1} \mid \mathbf{x}, T)] \mid \mathbf{x}]\}^{-1} \\
&= \{E[(1+T)/x_i \mid \mathbf{x}]\}^{-1} = \{E[(1+T) \mid \mathbf{x}]\}^{-1} x_i \qquad (1 \leqslant i \leqslant p).
\end{aligned} \tag{2.5}$$

Also, it follows from (2.4) that $\delta_i(\mathbf{x}) = 0$ if $x_i = 0$.

Now the joint pdf of $\mathbf{X}$ and T is given by

$$f(\mathbf{x}, t) = t^p (1+t)^{-(z+p)} g(t), \tag{2.6}$$

where $z = \sum_1^p x_i$. Thus, from (2.6) it follows that the conditional pdf of T given $\mathbf{X} = \mathbf{x}$ is given by

$$f(t \mid \mathbf{x}) \propto t^p(1+t)^{-(z+p)} g(t), \tag{2.7}$$

which depends on $\mathbf{x}$ only through z. Hence, $\{E[(1+T) \mid \mathbf{X}]\}^{-1} = \{E[(1+T) \mid Z]\}^{-1}$, and the proof of the lemma is complete.

Remark 1. A similar lemma for the normal case appeared in Strawderman [9], and later was used by Faith [5].

To prove the main result of this paper, we need next the following lemma due to Tsui and Press [13] generalizing Theorem 2.1 of Clevenson and Zidek [4].

LEMMA 2. *Let $\delta(\mathbf{x}) = (1 - \phi(z)/(z+a))\mathbf{x}$, where $p \geqslant 2$, $a > 0$, and $\phi(z)$ is a real-valued nondecreasing function of z with $\phi(z) \not\equiv 0$, and $0 \leqslant \phi(z) \leqslant 2\min(p-1, a)$. Then, the estimator $\delta(\mathbf{X})$ has risk $R(\boldsymbol{\theta}, \boldsymbol{\delta}) = \sum_1^p \theta_i^{-1} E_{\boldsymbol{\theta}}[\delta_i(\mathbf{X}) - \theta_i]^2$ less than p for all $\boldsymbol{\theta}$ such that $\theta_i \in (0, \infty)$, $1 \leqslant i \leqslant p$.*

The main result of the paper is as follows.

THEOREM 1. *Let the marginal pdf of T given by $g(t)$ satisfy $\lim_{t\to\infty} t^{1+\epsilon} g(t) < \infty$ for some $\varepsilon > 0$. If $g(t)$ is differentiable in t and if $h(t) = (g'(t)/g(t))t$ is well defined for $t > 0$, then the generalized Bayes estimator of $\boldsymbol{\theta}$ under L_1 in* (1.1) (*given in Lemma* 1) *has risk less than p if*

(i) *$h(t)$ is nonincreasing in t, and,*
(ii) $\lim_{t\to 0+} h(t) \leqslant 2a - (p+1)$, $0 < a \leqslant p-1$.

Proof. From Lemma 1, the generalized Bayes estimator of $\boldsymbol{\theta}$ is of the form $\boldsymbol{\delta}(\mathbf{X})$ given in Lemma 2 with

$$\begin{aligned} \phi(z) &= (z+a)\left(1 - \frac{1}{E(1+T \mid z)}\right) \\ &= (z+a)\frac{E(T \mid z)}{E(1+T \mid z)} \qquad \text{for} \quad z \geqslant 1, \end{aligned} \tag{2.8}$$

$$\phi(0) = 0. \tag{2.9}$$

Next note that from (2.7),

$$\phi(z) = (z+a)\frac{\int_0^\infty t^{p+1}(1+t)^{-(z+p)}\, g(t)\, dt}{\int_0^\infty t^p(1+t)^{-(z+p-1)}\, g(t)\, dt}. \tag{2.10}$$

Integrating the numerator by parts, and using $\lim_{t\to\infty} t^{1+\varepsilon}g(t) < \infty$ for some $\varepsilon > 0$, for $z \geqslant 1$,

$$\int_0^\infty t^{p+1}(1+t)^{-(z+p)}\, g(t)\, dt$$
$$= (z+p-1)^{-1}\left\{\int_0^\infty (p+1)\, t^p(1+t)^{-(z+p-1)}\, g(t)\, dt\right.$$
$$\left. + \int_0^\infty t^{p+1}(1+t)^{-(z+p-1)}\, g'(t)\, dt\right\}. \tag{2.11}$$

Hence, from (2.8), (2.10) and (2.11) for $z \geqslant 1$,

$$\phi(z) = (z+a)(z+p-1)^{-1}[(p+1) + E^*(h(T) \mid z)], \tag{2.12}$$

where E^* denotes the conditional expectation (given $Z = z$) with respect to a random variable T^* with conditional pdf given by

$$f^*(t \mid z) \propto t^p(1+t)^{-(z+p-1)}\, g(t), \qquad z \geqslant 1. \tag{2.13}$$

Now since $0 < a \leqslant p-1$, $(z+a)(z+p-1)^{-1}$ is $\uparrow$ in z. Also, from (2.13), for $0 < t_1 < t_2$, $1 \leqslant z_1 < z_2$,

$$f^*(t_1 \mid z_1) f^*(t_2 \mid z_2) - f^*(t_2 \mid z_1) f^*(t_1 \mid z_2)$$
$$= (t_1 t_2)^p\, g(t_1)\, g(t_2)[u(z_1)\, u(z_2)]^{-1}[(1+t_1)^{-(z_1+p-1)}(1+t_2)^{-(z_2+p-1)}$$
$$- (1+t_1)^{-(z_2+p-1)}(1+t_2)^{-(z_1+p-1)}], \tag{2.14}$$

where $u(z) = \int_0^\infty t^p(1+t)^{-(z+p-1)}\, g(t)\, dt$. Now

$$\frac{\{(1+t_1)^{z_1+p-1}(1+t_2)^{z_2+p-1}\}}{\{(1+t_1)^{z_2+p-1}(1+t_2)^{z_1+p-1}\}} = \left\{\frac{(1+t_2)}{(1+t_1)}\right\}^{z_2-z_1} > 1. \tag{2.15}$$

Hence,

$$\text{lhs of (2.14)} < 0 \text{ when } 0 < t_1 < t_2,\ 1 \leqslant z_1 < z_2. \tag{2.16}$$

This shows that the family of conditional pdf's $f^*(t \mid z)$ has monotone likelihood ratio in $-z$. Hence, since $h(t)$ is $\downarrow$ in t, using a well-known result (see Lemma 2 of Lehmann [7, p. 74]) it follows that

$$E^*[h(T) \mid z] \text{ is } \uparrow \text{ in } z \geqslant 1. \tag{2.17}$$

Hence, from (2.12) and (2.17), $\phi(z)$ is $\uparrow$ in z for $0 < a \leqslant p-1$. Also, since $P(T > 0) = 1$, $P(\phi(Z) > 0) > 0$.

Now using (ii) of Theorem 1,

$$\sup_{z>0} \phi(z) \leqslant p + 1 + 2a - (p+1) = 2a. \tag{2.18}$$

Since $0 < a \leqslant p - 1$, the conditions of Lemma 2 are satisfied, and hence it follows that the generalized Bayes estimator of $\boldsymbol{\theta}$, namely, $\hat{\boldsymbol{\delta}}(\mathbf{X})$ has risk less than p.

Remark 2. Since $\mathbf{X}$ is minimax under L_1, $\hat{\boldsymbol{\delta}}(\mathbf{X})$ is also minimax under L_1.

The following example illustrates the use of Theorem 1 in finding a class of proper Bayes minimax estimators of $\boldsymbol{\theta}$ for $p \geqslant 3$.

EXAMPLE. Let $g(t) \propto t^{m-1}(1+t)^{-(m+n)}$, $0 < m \leqslant p - 2$, $n > 0$, i.e., we consider a family of type II beta priors for T. Then,

$$h(t) = \left(\frac{g'(t)}{g(t)}\right) t = (m-1) - (m+n)\frac{t}{1+t}. \tag{2.19}$$

Hence, $h(t)$ is $\downarrow$ in t, $\lim_{t \to 0+} h(t) = m - 1$, $\lim_{t \to \infty} t^{1+\epsilon} g(t) < \infty$ by choosing $0 < \varepsilon \leqslant n$. Hence, from (2.7),

$$\begin{aligned} E(1+T \mid z) &= \frac{\int_0^\infty t^{m+p-1}(1+t)^{-(m+n+z+p-1)}\,dt}{\int_0^\infty t^{m+p-1}(1+t)^{-(m+n+z+p)}\,dt} \\ &= \frac{B(m+p, n+z-1)}{B(m+p, n+z)} \\ &= \frac{z+m+n+p-1}{z+n-1}, \qquad z \geqslant 1. \end{aligned} \tag{2.20}$$

Thus,

$$\phi(z) = (z+a)\left(1 - \frac{z+n-1}{z+m+n+p-1}\right) = \frac{(z+a)(m+p)}{z+m+n+p-1} \tag{2.21}$$

which is $\uparrow$ in z for $0 < a \leqslant m + n + p - 1$ and hence for $0 < a \leqslant p - 1$. Also, for $0 < a \leqslant m + n + p - 1$,

$$\sup_{z \geqslant 0} \phi(z) = m + p \leqslant 2(p-1) \qquad \text{since} \quad 0 < m \leqslant p - 2. \tag{2.22}$$

Hence, choosing $a \leqslant p - 1$,

$$\sup_{z \geqslant 0} \phi(z) \leqslant 2\min(p-1, a). \tag{2.23}$$

Thus, for the class of type II beta densities with $0 < m \leqslant p-2$, $n > 0$, $p \geqslant 3$, a class of proper Bayes minimax admissible estimates of $\boldsymbol{\theta}$ is given by

$$\boldsymbol{\delta}_{m,n}(\mathbf{x}) = \left(1 - \frac{m+p}{z+m+n+p-1}\right)\mathbf{x} = \frac{z+n-1}{z+m+n+p-1}\mathbf{x}. \quad (2.24)$$

Remark 3. Clevenson and Zidek [4] (see their Theorem 2.5) obtained a class of proper Bayes minimax estimates of $\boldsymbol{\theta}$ of the form

$$\tilde{\boldsymbol{\delta}}_{\beta}(\mathbf{x}) = (z/(z+\beta+p-1))\mathbf{x}, \qquad 1 < \beta \leqslant p-1, \quad p \geqslant 3. \quad (2.25)$$

Putting $m = \beta - 1$, $n = 1$ in (2.24), one finds that the class of estimators (2.25) is a subclass of estimators (2.24). It should be noted that Clevenson and Zidek [4] took an apparently different approach of first reparametrizing $(\theta_1, ..., \theta_p)$ into $(u, v_1, ..., v_{p-1})$, where $u = \sum_1^p \theta_i$ and $v_i = \theta_i/u$ $(1 \leqslant i \leqslant p-1)$, and then putting a prior of the form $m_\beta(u)\, g(v_1, ..., v_{p-1})$ to u and $(v_1, ..., v_{p-1})$, where g is Dirichlet(1,..., 1; 1) (see Wilks [14, p. 177] for the definition of a Dirichlet distribution), and

$$m_\beta(u) \propto \int_0^\infty (1+ut)^{-\beta} \exp(-t^{-1})\, t^{-p}\, dt$$

$$= \int_0^\infty (1+uz^{-1})^{-\beta} \exp(-z)\, z^{p-2}\, dz, \quad (2.26)$$

putting $z = t^{-1}$.

It is anticipated from (2.24) and (2.25) that the prior used by Clevenson and Zidek [4] is a special case of our two stage prior. This is, in fact, true and is now exhibited in details.

First note that conditional on $T = t$, θ_i's are iid exponential with parameter t. Hence, conditional on $T = t$, u is distributed independently of $(v_1, ..., v_{p-1})$. Further, the conditional pdf of u given $T = t$ is

$$f(u \mid t) \propto t^p \exp(-tu)\, u^{p-1}, \qquad u > 0, \quad t > 0, \quad (2.27)$$

which is a gamma (t, p) density, while conditional on $T = t$, $(v_1, ..., v_{p-1})$ has a Dirichlet(1,..., 1; 1) distribution. Thus, marginally $(v_1, ..., v_{p-1}) \sim$ Dirichlet(1,..., 1; 1), while the marginal pdf of u is

$$\int f(u \mid t)\, g(t)\, dt. \quad (2.28)$$

Putting $g(t) \propto t^{\beta-2}(1+t)^{-\beta}$, $1 < \beta \leqslant p-1$, it follows from (2.27) and (2.28) that the marginal pdf of u is given by

$$m_\beta^0(u) \propto \int_0^\infty \exp(-tu)\, t^{p+\beta-2} u^{p-1}(1+t)^{-\beta}\, dt$$

$$= \int_0^\infty \exp(-z)(z/u)^{p+\beta-2}\, u^{p-1}(1+zu^{-1})^{-\beta}\, u^{-1}\, dz$$

$$= \int_0^\infty \exp(-z)\, z^{p+\beta-2}(z+u)^{-\beta}\, dz$$

$$= \int_0^\infty (1+uz^{-1})^{-\beta} \exp(-z)\, z^{p-2}\, dz, \tag{2.29}$$

which is the same as $m_\beta(u)$.

Remark 4. Next we show that conditions (i) and (ii) in Theorem 1 are not met by any proper prior when $p=2$. First note that when $p=2$, from (ii),

$$\lim_{t\to 0+} h(t) \leqslant 2\min(1, a) - 3 \leqslant -1. \tag{2.30}$$

Hence, since $h(t)$ is $\downarrow$ in t,

$$g'(t)/g(t) \leqslant -t^{-1} \qquad \text{for all } t > 0. \tag{2.31}$$

Integrating both sides of (2.31) over the interval $[t_1, t_2]$ where $0 < t_1 < t_2$ one gets

$$\log[g(t_2)/g(t_1)] \leqslant \log(t_1/t_2). \tag{2.32}$$

Hence, $g(t_1) \geqslant g(t_2)\, t_2/t_1 = c/t_1$ (say) for fixed t_2 which is not integrable between 0 and t_2.

Remark 5. Strawderman [10] has shown in the multinormal case that for $p=3$ or 4, there does not exist any spherically symmetric proper Bayes minimax estimator of $\boldsymbol{\theta}$. It is worth investigating whether or not in the Poisson case there exists any proper Bayes minimax estimate of the form $u(z)\mathbf{x}$ for $\boldsymbol{\theta}$ in the case $p=2$.

Remark 6. We do not know whether there exist any proper Bayes minimax estimator $\mathbf{X}$ under the loss L_k in (1.1) with $k \neq 1$. Proper Bayes estimators of $\boldsymbol{\theta}$ can be generated in such cases by using two stage priors, the first stage prior being a gamma. Such estimators, however, do not belong to the class given in Theorem 2 of Tsui and Press [13], and it is not known as yet whether they dominate $\mathbf{X}$.

References

[1] Berger, J. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of Gamma scale parameters. *Ann. Statist.* **8** 545–571.

[2] Brown, L. D. (1980). Examples of Berger's phenomenon in the estimation of independent normal means. *Ann. Statist.* **8**, 572–585.

[3] Brown, L. (1979). A heuristic method for determining admissibility of estimators with application. *Ann. Statist.* **7** 960–994.

[4] Clevenson, M. L. and Zidek, J. V. (1975). Simultaneous estimation of the means of independent Poisson laws. *J. Amer. Statist. Assoc.* **70** 698–705.

[5] Faith, R. E. (1978). Minimax estimators of a multivariate normal mean. *J. Multivariate Anal.* **8** 372–379.

[6] Hudson, H. M. (1978). A natural identity for exponential families with applications in multiparameter estimation. *Ann. Statist.* **6** 473–484.

[7] Lehmann, E. L. (1959). *Testing Statistical Hypothesis.* Wiley, New York.

[8] Peng, J. C. (1975). *Simultaneous Estimation of the Parameters of Independent Poisson Distribution.* Technical Report No. 78, Department of Statistics, Stanford University, Stanford, Calif.

[9] Strawderman, W. E. (1971). Proper Bayes minimax estimators for the mean of a multivariate normal population. *Ann. Math. Statist.* **42** 385–388.

[10] Strawderman, W. E. (1972). On the existence of proper Bayes minimax estimators of the mean of a multivariate distribution. In *Proc. 6th Berk. Symp. Math. Stat. Prob.*, Vol. VI, pp. 51–55. Univ. of California Press, Berkley.

[11] Tsui, K. W. (1978a). Multiparameter estimation of discrete exponential distributions. *Canad. J. Statist.*, in press.

[12] Tsui, K. W. (1978b). *Simultaneous Estimation of Several Poisson Parameters under Squared Error Loss.* Technical Report No. 37, Department of Statistics, University of California, Riverside.

[13] Tsui, K. W. and Press, S. J. (1978). *Simultaneous Estimation of Several Poisson Parameters under* k-*Normalized Squared Error Loss.* Technical Report No. 38, Department of Statistics, University of California, Riverside.

[14] Wilks, S. S. (1962). *Mathematical Statistics.* Wiley, New York.

Printed in Belgium